



**Sentiment Urgency Emotion Conversion Over Time for
Business Intelligence**

Journal:	<i>International Journal of Web Information Systems</i>
Manuscript ID	Draft
Manuscript Type:	Research Paper
Keywords:	Social media mining, sentiment analysis, opinion mining, urgency detection, sentiment conversion

SCHOLARONE™
Manuscripts

Sentiment Urgency Emotion Conversion Over Time for Business Intelligence

Abstract. Social media has become a vital part of any institute’s marketing plan. Social networks benefit businesses by allowing them to interact with their clients, grow brand exposure through offers and promotions, and find new leads. It also offers vital information concerning the general emotions and sentiments directly connected to the welfare and security of the online community involved with the brand. Big organizations can make use of their social media data in order to generate planned and operational decisions. In this work, a model called Sentiment Urgency Emotion Detection (SUED) from previous work will be applied on tweets from two different periods of time, one before the start of the COVID-19 pandemic and the other after it started in order to monitor the conversion of sentiments and emotions over time. The model has been trained to improve its accuracy and F1 score so that the precision and percentage of correctly predicted texts is high. This model will be tuned to improve results [15] [16] and will be applied on a general business Twitter account of one of the largest chains of supermarkets in the United Kingdom to be able to see what sentiments and emotions can be detected, and how urgent they are. This will show the effect of COVID-19 pandemic on the conversions of the sentiments, emotions, and urgencies of tweets.

1 Introduction

Social media is categorized as a group of online platforms, websites, or applications that permit users to interact, interconnect, cooperate, and discuss point of views and information [3][4][19]. It can also be regarded as a two-way platform

where users can share and discuss information [7]. Social networking websites have developed into a source of information since people can post real time posts concerning their opinions on topics and present issues or display negative comments or complaints [1]. It has unraveled innovative capacities for businesses to connect with their customers and employees by providing them with the option of sending messages quickly and obtaining real-time responses [9]. This has enabled users to connect and to expand their marketing opportunities through a growth of communication. The democratization of information has permitted companies to communicate with customers and for customers to communicate with each other making this model a “two-way” communication model [10].

Some studies have focused on the use of Twitter in businesses. A study has shown that the Twitter account for an organization was used for a mixture of client testaments, dissatisfaction, replies to any enquiries or complaints, and questions and answers [5]. It was also seen that firms were unlikely to post any type of “me now” tweets on their Twitter accounts but were more likely to post informative posts that encourage followers to be involved with online conversation. Therefore, this study showed that Twitter can be considered a good means for client relationship administration even if the firms were not using Twitter for this reason [5].

Another research was also done on five hundred Fortune companies but only took 93 active companies as a sample for comprehending the interaction characteristics of Twitter and the public groups targeted [14]. The research result from examining their tweets showed that the percentage of replies to posts from users was 60.2%, the percentage of sharing important data about the company was 58.1%, and the percentage of making enquiries was 30.1%. The results also showed that the tweets targeted for all followers in general was 74.5% and was greater than the tweets for specific users which was 0.9% [14].

Some work focused on the word-of-mouth discussion on Twitter and the results showed that about one fifth of all tweets hold the name of a brand, good, or facility [5]. In addition, one fifth of these word-of-mouth tweets showed few sentiments. Also, the study showed that the positive tweets are more than half of the branded tweets while the negative tweets are only one third of them [5]. This study concluded that the linguistic assembly of tweets is related to the linguistic patterns of natural language expressions and it showed that Twitter is a rich word-of-mouth site for institutes to explore as part of their global branding policy [5].

Another research was performed to recognize how firms communicate with their employees over Twitter. 5245 business accounts from these firms were taken into consideration for the study during which 1000 Tweets were examined and classified according to five classes. The classes were happiness with some reference, specific engagement, alerts about products, broadcasts, and engagement. The broadcasts class was divided into two classes itself, which are deals and events. The deal class showed that it was of 95% positive predictive value and 93% sensitivity, while the events class showed that it was of 96% positive predictive value and 97% sensitivity [13].

Twitter is a good platform to extract data for text categorization. Automatic text classification is hard for a computer due to the huge number of words. The categorization of tweets is even more difficult due to the restriction on the number of words in a tweet [6]. Classifying emotions in text is very hard for two reasons. First, feelings can be implied and activated by precise actions or circumstances. Second, gathering different emotions based on keywords can be very difficult to detect [6] [18]. Classifying urgency in text might not be easy. Previous work shows a social network-based solution that can observe multiple social networks to detect keywords, urgency ratings, the request owner's identity, date, and time [2]. It also governs which posts or chats are crucial and prioritize or rank pending urgent concerns [2].

In this work, SUEd model will be applied on tweets of one of the largest chains of supermarkets in the United Kingdom. The tweets will be extracted from two different periods such that the first period is before the start of the COVID-19 pandemic [15], and the other one is after the pandemic started. This work will discover if the tweets have a sentiment, will realize if there are emotions in them, and will detect if they are urgent. Therefore, this will assist in determining the gratification or the dissatisfaction of online communities around the brand being examined. Moreover, sentiments will be compared between the two periods in order to evaluate how sentiments and emotions vary over time taking into consideration the COVID-19 as an affective factor. In addition, SUEd will be tuned to enhance results and the knowledge that is

mined when turning data into decisions is crucial since it will aid stakeholders handling the institute to evaluate the topics and issues that were mostly emphasized

2 Methods

In this work, the SUED model from previous work [15] will be applied based on several classifiers using monkey learn platform [11]. The classifiers combined are sentiment analysis, urgency detection, and emotion classification. The model is built using Support Vector Machines Algorithm (SVM) which is regarded as a controlled learning method [12] [17]. It shows input-output mapping functions from a set of labeled training data [17] and is based on the statistical learning method [12]. If it's a classification function, data introduced is converted using nonlinear kernel functions to a high-dimensional feature space to be made more separable in contrast to the first input space [17]. If the function is regression, it disregards any training data that is adequately adjacent to the model prediction [17].

The N-gram range provides the characteristics category to be executed to characterize texts [11]. Applying the SUED model from previous work [15][16], the N-gram applied for this model is Unigrams or words (n-gram size = 1) and Bigrams or terms compounded by two words (n-gram size = 2) [11]. The Max features for this model is 10,000 which adjusts the maximum number of features to be used to characterize texts in the training/classification process and impacts the computation resources required to train the model [11]. In order to enhance the model, SUED will be tuned to create another model SUED2 such that the N-gram applied for this model is Unigrams, Bigrams, and Trigrams or terms compounded by up to three words (n-gram size = 3) [11]. In addition, the Max features will be doubled to 20,000 to improve results.

The categories implemented for this work and previous work [15] [16] are associated with the three classifiers and are revealed in Table 1. This model was trained through a learning mechanism on a training data set containing tweets from multiple sources in order to improve its accuracy and F1 score [11] [15] [16]. It learns to cluster the tweets into specific categories by matching input or keywords to respective matching output or categories.

SUED model will be applied on tweets from two different periods T1 and T2 related to the same supermarket chain store as a business entity, T1 before the start of the COVID-19 pandemic [15], and T2 other after it. This will help identify the categories of the feedback from the customers of this chain which can be customer service, reviews, product reviews, or suggestions to improve their service. At the same time, the model can categorize the chain's reply feedbacks that either the customer service provided, or other individuals provided. Moreover, it will try to categorize the texts if they are urgent and require attention and it will provide confidence numbers for these categories. SUED2 will also be applied on period T2 to monitor enhancement of the results.

The same experimental environment will be used for this work as previous work [15] [16] which is a Windows 10 Enterprise laptop with a processor of Intel® Core™

i5-8250U CPU @ 1.60GHz 1.80GHz. The installed memory (RAM) is 8.00 GB. The System type is 64-bit Operating System, x64-based processor.

Two data sets were used for this work. The first data set is of data size 2795 recent tweets where the store chain brand has been mentioned before the start of the COVID-19 pandemic up until 17th of February 2020. The second data set is of data size 2795 recent tweets where the store chain brand has been mentioned after the start of the COVID-19 pandemic up until 12th of June 2020.

Like previous work [15] [16], the model was fed with training data to enhance its performance. The machine learning algorithm learns through tagging and retagging of the sample tweets of the training data that there is a specific category or categories expected as a result of an input have specific keywords [11]. Enhancing the model requires improving two substantial metrics which accuracy and F1 Score (precision and recall). The accuracy is the percentage of test tweets that were matched with the correct category. It is calculated as the quotient of the correctly classified tweets by the total number of tweets in test data set. F1 score on the other hand merges both precision and recall [8] [11]. Recall is the proportion of positive sentiments which are correctly acknowledged while precision is the ratio between the correct sentiments predicted to the total number of matches predicted [8]. Fig. 1. shows the accuracy and F1 score following the training of the model. Fig. 2. shows an overview of the most related keywords as the result of the training [11][15].

3 Results

Following the training of the SUED model, the model was run on the same data set from previous work [15] over the time period T1 before the COVID-19 pandemic, and it was run on another data set over the time period T2 after the COVID-19 pandemic. Two batches for T1 and T2 were imported to SUED each having 2795 tweets. Processing the datasets will provide the category or categories for each tweet and its/their respective confidence value(s). 151 and 155 single or combined categories have been produced for T1 and T2 respectively for SUED. 162 single or combined categories have been produced from T2 for SUED2. Fig. 3. and Fig. 4. show the top 20 categories containing the largest group of tweets for T1 and T2 for SUED. Fig. 7. shows the top 20 categories containing the largest group of tweets for T2 for SUED2. Fig.5. and Fig.6. display the top 10 categories percentages out of the overall for T1 and T2 respectively. Similarly, Fig.8. displays the top 10 categories percentages out of the overall for T2 for SUED2.

For each tweet in T1 and T2, the model calculates the confidence value for every single category that matches it. Afterwards, the average confidence

values of all the categories for each tweet is then calculated. Based on this, the average confidence value of each of the top 10 categories from Fig.3., Fig.4, and Fig.7. can be calculated. It will be the average of the average confidences for all the tweets belonging to the same category. The average confidence of the top 10 categories for T1 and T2 for SUEd is observed in Table 2 and Table 3 respectively. The average confidence of the top 10 categories for T2 for SUEd2 is observed in Table 4.

4 Discussion

Upon analyzing the results, most of the tweets have been classified as "Feedback" category. For SUEd, the number of tweets which were categorized as "Feedback and Positive" have decreased by 15.57% while the number of tweets which were categorized as "Feedback and Complaint" have increased by 37.13%. Thus, the sentiments of users around this brand have converted over time more towards negativity. One of the main factors for this can be the COVID-19 pandemic. Comparing Table 3 and Table 4 shows that upon the increase of the Max features, we were able to get higher confidence numbers by 6 out of 10 of the top 10 categories for T2. Since many retail stores have temporarily closed during the pandemic, most customers have shifted towards online shopping and making their orders through e-commerce. As a result, more orders have been made online, which might have been more than the institute can accommodate. With the quarantine and huge number of orders, these orders might not have been received on their expected delivery dates, thus customers were not satisfied, and more complaints would have been tweeted, with less compliments.

Customer Service for this business can benefit from this model since it provides them the knowledge about the classifications of the tweets, the urgent complaints that might need assistance, the requests coming from customers, and the feedback.

Some difficulties were noted. In Fig. 3. and Fig. 4., 93 tweets in T1 and 69 tweets in T2 which constitute 3.33% and 2.47% of their respective entire datasets could not be categorized to any of the categories of SUEd which might be due to the model not being able to match the keywords in the tweets to any of the groups. Other challenges faced were tweets that may have not been properly grouped due to false positive or

false negative in the tweets. Another problem faced as well was that 29 tweets in T1 and 34 tweets in T2 which constitute around 1% of their datasets showed irrelevant confidence numbers.

5 Conclusion

In this work, a learning model for sentiment, urgency, and emotion detection (SUED) has been applied for social media mining and opinion mining. The model was trained on text samples from many twitter sources to improve its accuracy and F1 Score. The model was run on a general business Twitter account for a supermarket chain on two different periods to measure the confidence of its categorization and to monitor the conversion of sentiments over time taking COVID-19 pandemic as a factor. In addition, SUED was tuned to create SUED2 to improve results. Tweets were able to be classified into one or multiple categories. Results showed that the sentiments of users around this brand has converted over time more towards negativity. Future work can involve training the model on a larger dataset so that it has a larger accuracy and F1 Score. This can help result in more correct classification for tweets and less unclassified tweets. This can also help get more relevant keywords to better associate categories to tweets. Future work can also involve focusing on tweet replies from institutes to monitor how customer service deals with feedback.

6 References

1. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. J. (2011, June). Sentiment analysis of twitter data. In *Proceedings of the workshop on language in social media (LSM 2011)* (pp. 30-38).
2. Chavez, D. L., Mohler, D. S., & Shockley, B. A. (2013). *U.S. Patent No. 8,515,049*. Washington, DC: U.S. Patent and Trademark Office.
3. Doyle, C. (2010). A literature review on the topic of social media, file retrieved from blog post.

4. Holotescu, C., & Grosseck, G. (2013). An empirical analysis of the educational effects of social media in universities and colleges. *Internet Learning*, 2(1), 5.
5. Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, 60(11), 2169-2188.
6. Janssens, O., Slembrouck, M., Verstockt, S., Van Hoecke, S., & Van de Walle, R. (2013, August). Real-time emotion classification of tweets. In *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)* (pp. 1430-1431). IEEE.
7. Johnston, R. (2011). Social media strategy: follow the 6 P's for successful outreach. *Alaska Business Monthly*, 27(12), 83-85.
8. Kim, Y., Jeong, S. R., & Ghani, I. (2014). Text opinion mining to analyze news for stock market prediction. *Int. J. Advance. Soft Comput. Appl*, 6(1), 2074-8523.
9. Loveiov, K., Waters, R. D., & Saxton, G. D. (2012). Engaging stakeholders through Twitter: How nonprofit organizations are getting more out of 140 characters or less. *Public relations review*, 38(2), 313-318.
10. Markos-Kujbus, É. and Gati, M., 2013. Social media's new role in marketing communication and its opportunities in online strategy building. *Institute of Marketing and Media*, pp.1-17.
11. Monkey Learn. (2013). Retrieved from <http://www.monkeylearn.com>
12. Pham, B. T., Prakash, I., Khosravi, K., Chapi, K., Trinh, P. T., Ngo, T. Q., ... & Bui, D. T. (2019). A comparison of Support Vector Machines and Bayesian algorithms for landslide susceptibility modelling. *Geocarto International*, 34(13), 1385-1407.
13. Popescu, A.M. and Jain, A., 2011, March. Understanding the functions of business accounts on twitter. In *Proceedings of the 20th international conference companion on World wide web* (pp. 107-108). ACM.
14. Rybalko, S. and Seltzer, T., 2010. Dialogic communication in 140 characters or less: How Fortune 500 companies engage stakeholders using Twitter. *Public Relations Review*, 36(4), pp.336-341.
15. Soussan, T., & Trovati, M. (2020). Sentiment Urgency Emotion Detection for Business Intelligence. In *Research Perspectives in Data Science and Smart Technology for Shipping Industries*.
16. Soussan, T., & Trovati, M. (2020, September). Improved Sentiment Urgency Emotion Detection for Business Intelligence. In *International Conference on Intelligent Networking and Collaborative Systems*. Springer, Cham.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

17. Wang, L. (Ed.). (2005). *Support vector machines: theory and applications* (Vol. 177). Springer Science & Business Media.

18. Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. (2012, September). Harnessing twitter" big data" for automatic emotion identification. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing* (pp. 587-592). IEEE.

19. Zeng, L., Hall, H. and Pitts, M.J., 2012. Cultivating a community of learners: The potential challenges of social media in Higher Education. *Social media: Usage and impact*, pp.111-126.

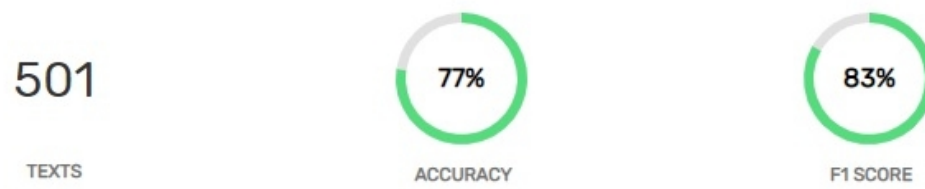


Fig. 1. Accuracy and F1 Score metrics for training the SUED Model

160x36mm (96 x 96 DPI)

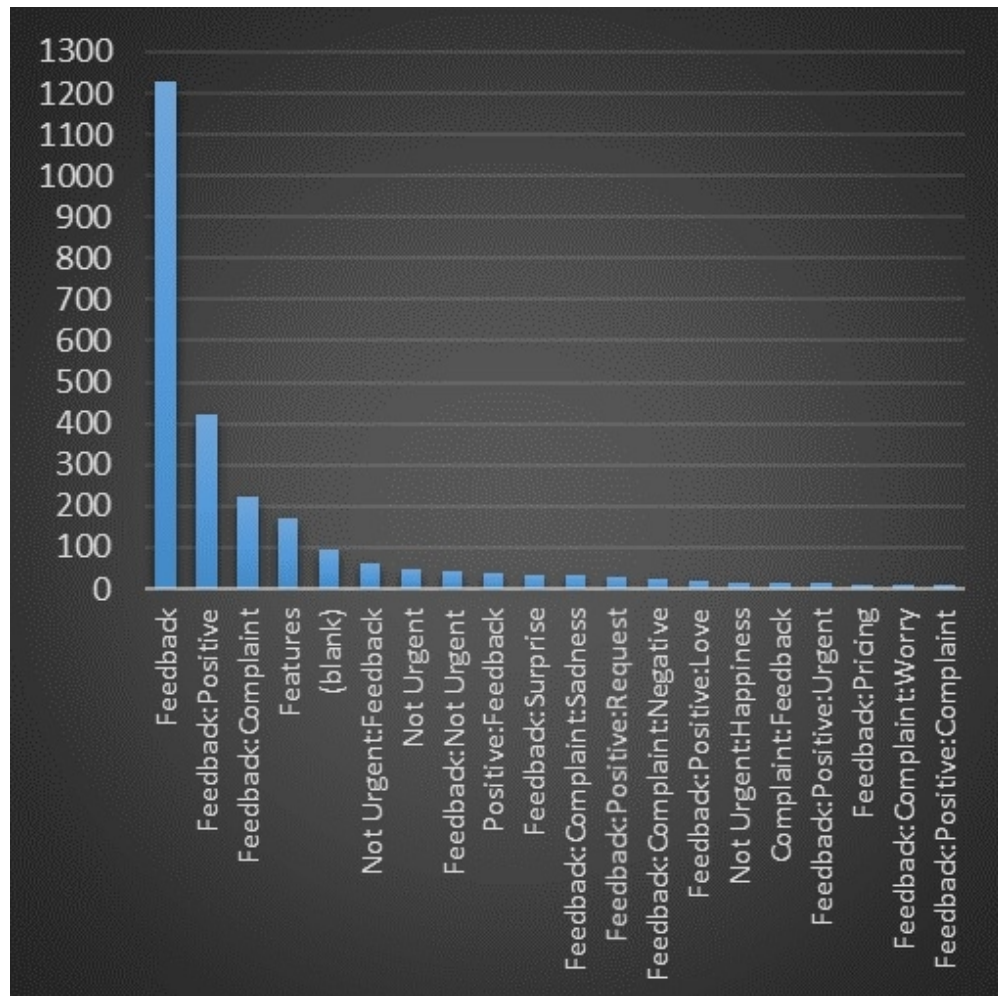


Fig. 3. Top 20 Categories having the highest number of tweets for T1 for SUEd

135x134mm (96 x 96 DPI)

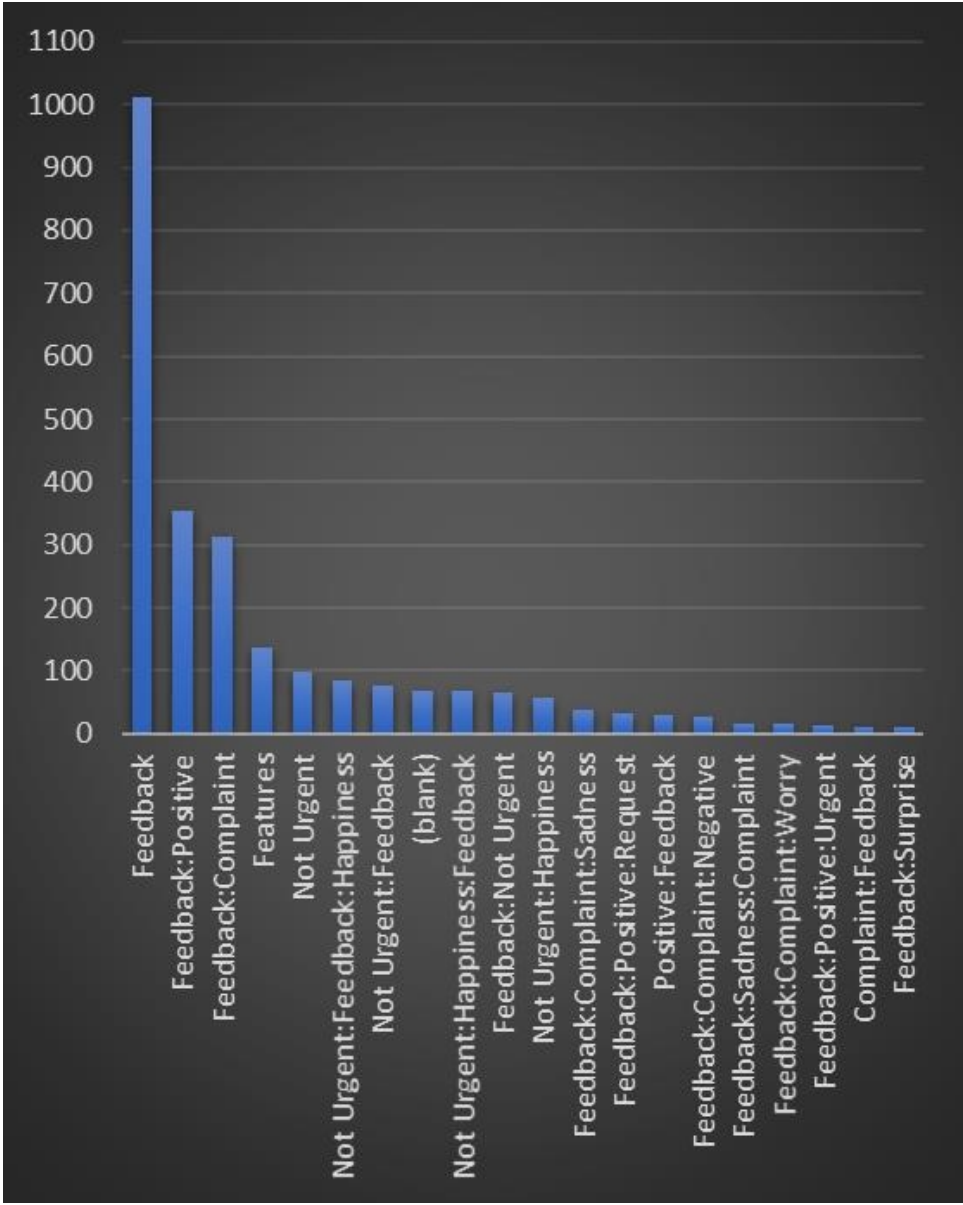


Fig. 4. Top 20 Categories having the highest number of tweets for T2 for SUEd

147x183mm (96 x 96 DPI)

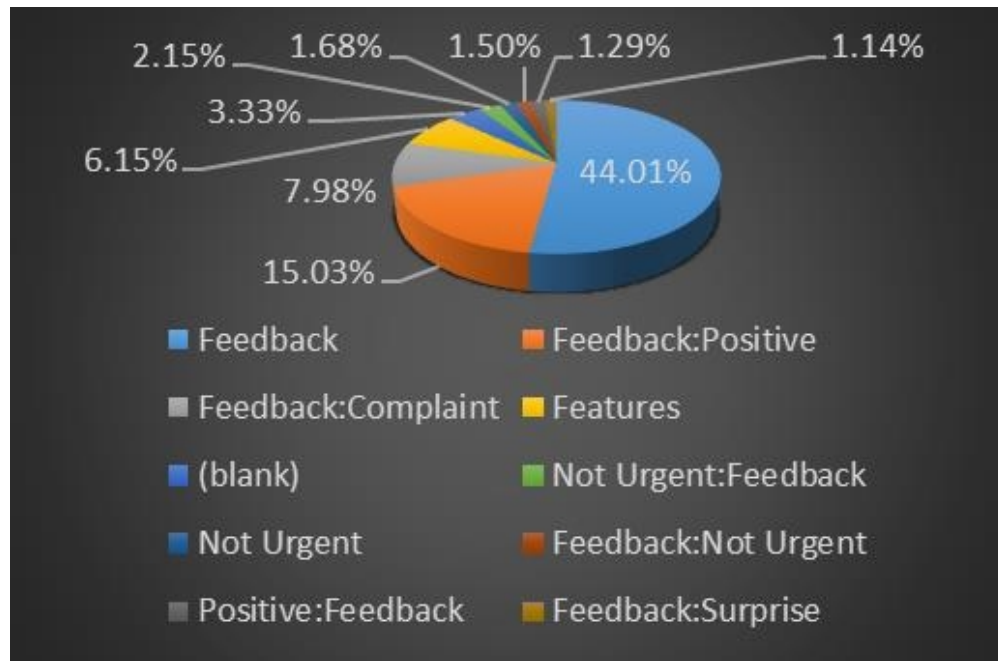


Fig. 5. Top 10 Categories percentages for T1 for SUEd

136x89mm (96 x 96 DPI)

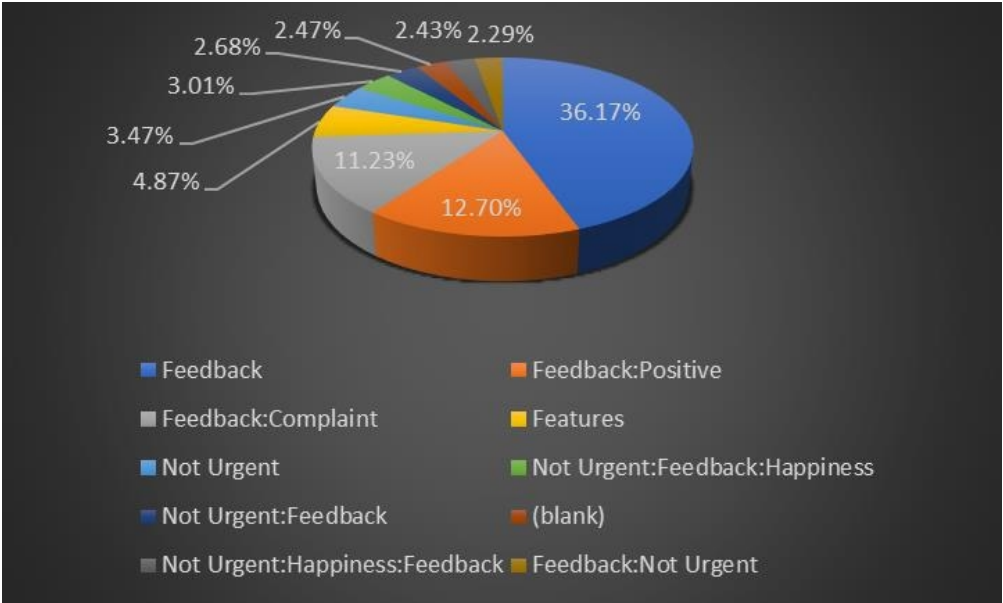


Fig. 6. Top 10 Categories percentages for T2 for SUED

190x114mm (96 x 96 DPI)

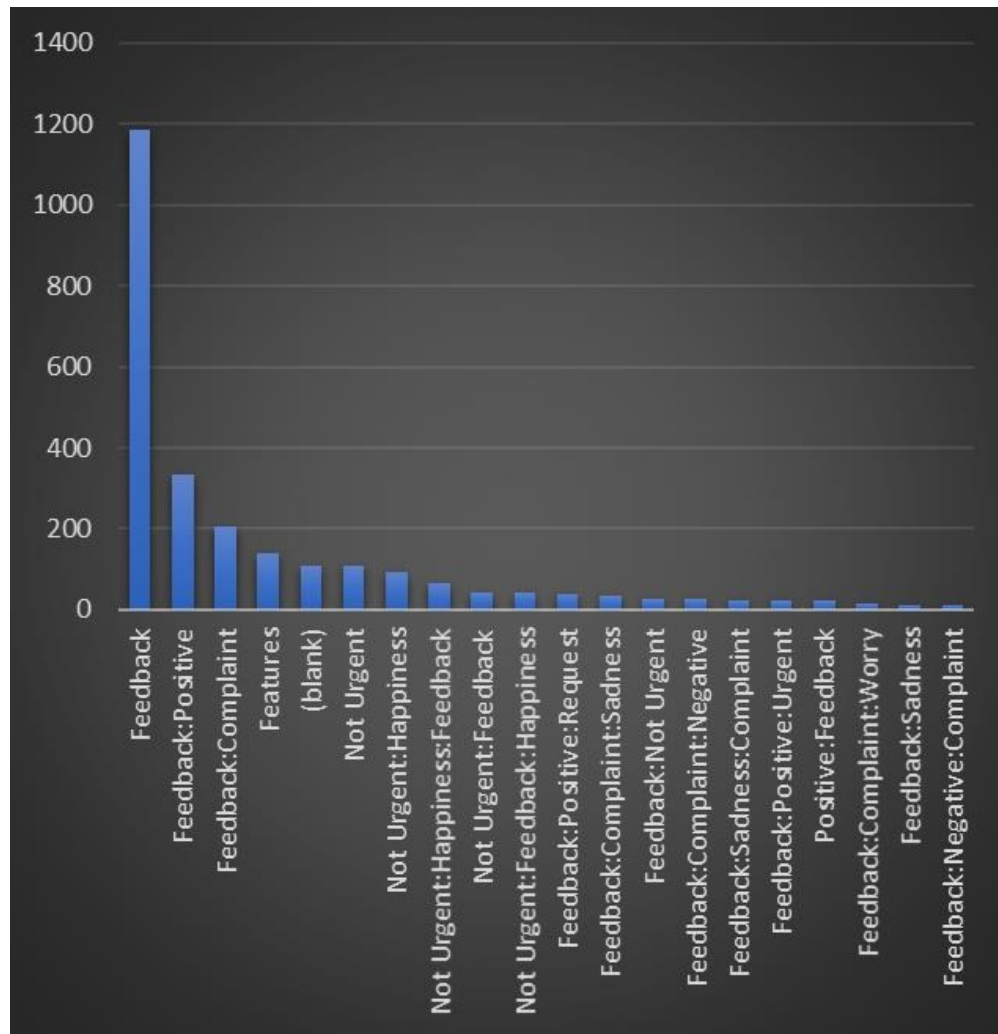


Fig. 7. Top 20 Categories having the highest number of tweets for T2 for SUE2

168x173mm (96 x 96 DPI)

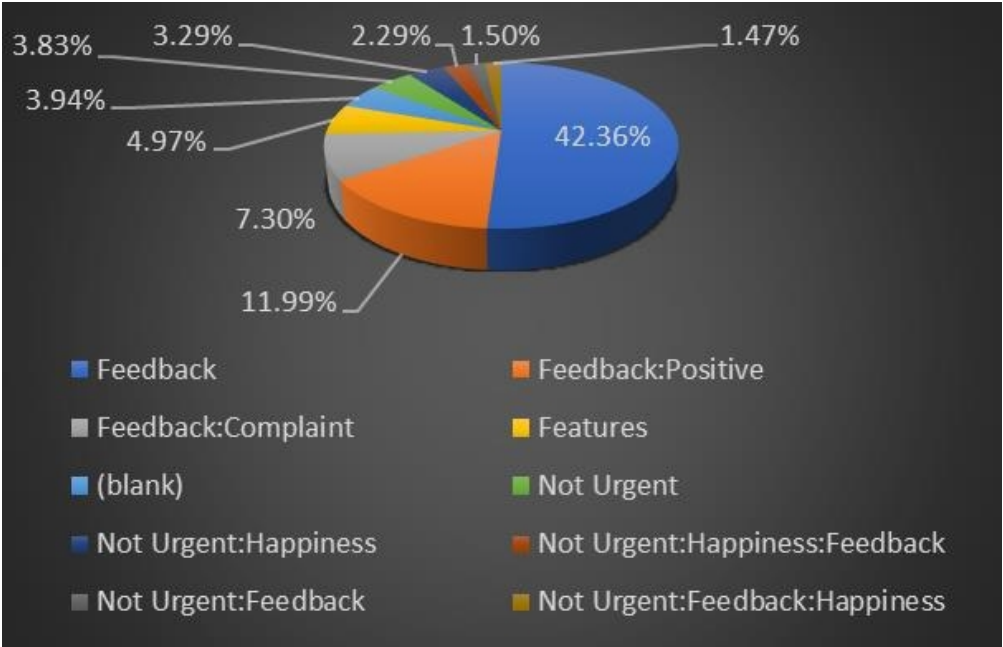


Fig. 8. Top 10 Categories percentages for T2 for SUEd2

160x103mm (96 x 96 DPI)

CATEGORIES

Anger
Features
Complaint
Positive
Negative
Pricing
Feedback
Request
Happiness
Sadness
Hate
Surprise
Love
Worry

TABLE 1 CATEGORIES FOR THE SUED MODEL

134x107mm (96 x 96 DPI)

Categories	AVERAGE CONFIDENCE
Feedback	0.752
Feedback: Positive	0.735
Feedback: Complaint	0.718
Features	0.685
Not Urgent: Feedback	0.657
Positive: Feedback	0.748
Feedback: Surprise	0.741
Feedback: Complaint: Sadness	0.780
Feedback: Positive: Request	0.814
Feedback: Complaint: Negative	0.796

TABLE 2 AVERAGE CONFIDENCE OF THE TOP 10 CATEGORIES FOR T1 FOR SUED

141x100mm (96 x 96 DPI)

Categories	AVERAGE CONFIDENCE
Feedback	0.765
Feedback: Positive	0.738
Feedback: Complaint	0.730
Features	0.650
Not Urgent: Feedback	0.674
Not Urgent	0.720
Not Urgent: Feedback:Happiness	0.684
Not Urgent:Happiness:Feedback	0.675
Feedback: Not Urgent	0.653
Not Urgent: Happiness	0.739

TABLE 3 AVERAGE CONFIDENCE OF THE TOP 10 CATEGORIES FOR T2 FOR SUED

141x101mm (96 x 96 DPI)

Categories	AVERAGE CONFIDENCE
Feedback	0.774
Feedback: Positive	0.745
Feedback: Complaint	0.724
Features	0.659
Not Urgent: Feedback	0.641
Not Urgent	0.724
Not Urgent: Feedback:Happiness	0.658
Not Urgent:Happiness:Feedback	0.676
Feedback:Positive:Request	0.773
Not Urgent: Happiness	0.743

TABLE 4 AVERAGE CONFIDENCE OF THE TOP 10 CATEGORIES FOR T2 FOR SUED2

141x100mm (96 x 96 DPI)